Development of genomic array footprinting for identification of conditionally essential genes in Streptococcus pneumoniae

Bijlsma, Jetta J. E.; Burghout, Peter; Moosterman, Tomas G.; Bootsma, Flester J.; de Jong, Anne; Hermans, Peter W. M.; Kuipers, Oscar P.; Kloosterman, Tomas G.; Bootsma, Hester J.

Published in:
Applied and environmental microbiology

DOI:
10.1128/AEM.01900-06

Publication date:
2007

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Mar. 2007, p. 1514–1524 Vol. 73, No. 5
0099-2240/07/$08.00+0 doi:10.1128/AEM.01900-06

Development of Genomic Array Footprinting for Identification of Conditionally Essential Genes in *Streptococcus pneumoniae*▽†

Jetta J. E. Bijlsma,[1]‡ Peter Burghout,[2]‡ Tomas G. Kloosterman,[1]‡ Hester J. Bootsma,[2]‡ Anne de Jong,[1] Peter W. M. Hermans,[2] and Oscar P. Kuipers[1]*

Department of Molecular Genetics, University of Groningen, Haren, The Netherlands,[1] and Laboratory of Pediatric Infectious Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands[2]

Received 9 August 2006/Accepted 30 December 2006

***Streptococcus pneumoniae* is a major cause of serious infections such as pneumonia and meningitis in both children and adults worldwide. Here, we describe the development of a high-throughput, genome-wide technique, genomic array footprinting (GAF), for the identification of genes essential for this bacterium at various stages during infection. GAF enables negative screens by means of a combination of transposon mutagenesis and microarray technology for the detection of transposon insertion sites. We tested several methods for the identification of transposon insertion sites and found that amplification of DNA adjacent to the insertion site by PCR resulted in nonreproducible results, even when combined with an adapter. However, restriction of genomic DNA followed directly by in vitro transcription circumvented these problems. Analysis of parallel reactions generated with this method on a large *mariner* transposon library showed that it was highly reproducible and correctly identified essential genes. Comparison of a *mariner* library to one generated with the in vivo transposition plasmid pGh:IS*S1* showed that both have an equal degree of saturation but that 9% of the genome is preferentially mutated by either one. The usefulness of GAF was demonstrated in a screen for genes essential for surviving zinc stress. This identified a gene encoding a putative cation efflux transporter, and its deletion resulted in an inability to grow under high-zinc conditions. In conclusion, we developed a fast, versatile, specific, and high-throughput method for the identification of conditionally essential genes in *S. pneumoniae*.**

Infection with the gram-positive pathogen *Streptococcus pneumoniae* is a worldwide cause of mortality and morbidity. Carriage of the bacterium can be asymptomatic but often progresses to diseases such as sinusitis and otitis media and to more serious infections such as pneumonia, sepsis, and meningitis. It is estimated that over a million people die of *S. pneumoniae*-related diseases every year (4, 27). Current polysaccharide vaccines are effective, but they only protect against a fraction of the 90 serotypes known, and replacement and disease by nonvaccine serotypes is already being observed (34). Treatment of *S. pneumoniae* infections is also confounded by the rise of strains resistant to the most commonly used antibiotics (19). Thus, there is an urgent need for the identification of new protein leads for the development of vaccines and antimicrobial drugs, preferably by using high-throughput, genome-wide screening methods.

Several methods have been used to determine which genes are needed by *S. pneumoniae* in the various niches it occupies in the host (conditionally essential genes), such as transcriptome analysis (29), differential fluorescence induction (22), and signature-tagged mutagenesis (STM) (7, 16, 30). Of these methods, STM is the only one that enables negative screening and thus directly addresses the essentiality of a gene under a certain condition. However, STM is hampered by various drawbacks: every mutant has to be grown and stored separately, which makes it hard to generate large libraries with a more than 1× coverage of the genome; detection of the tags is cumbersome; and identification of transposon insertion sites is labor intensive (2). In the three in vivo STM screens performed with *S. pneumoniae*, no more than 8,000 mutants in total have been tested. Furthermore, there is little overlap in the identified genes in each study, indicating that the screens were not saturated, and more importantly, that not all existing conditionally essential genes have been identified (7, 16, 30).

Therefore, we set out to develop genomic array footprinting (GAF), a high-throughput method to identify conditionally essential genes in *S. pneumoniae* by using a combination of random transposon mutagenesis and microarray technology (Fig. 1). GAF detects the transposon insertion sites in a library by amplifying and labeling the chromosomal DNA adjacent to the transposon and subsequent hybridization of these probes to a microarray. Identification of transposon insertion sites in mutants that have disappeared from the library due to selection, which represent conditionally essential genes, is achieved by differential hybridization of the probes generated from the library grown under two conditions to an array (Fig. 1). We anticipated that the most critical step in the whole procedure is the specific amplification and detection of the DNA adjacent

* Corresponding author. Mailing address: Department of Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Rijksuniversiteit Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands. Phone: 31-50-3632093. Fax: 31-50-3632348. E-mail: o.p.kuipers@rug.nl.

‡ These authors contributed equally to the work described in this paper.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 19 January 2007.

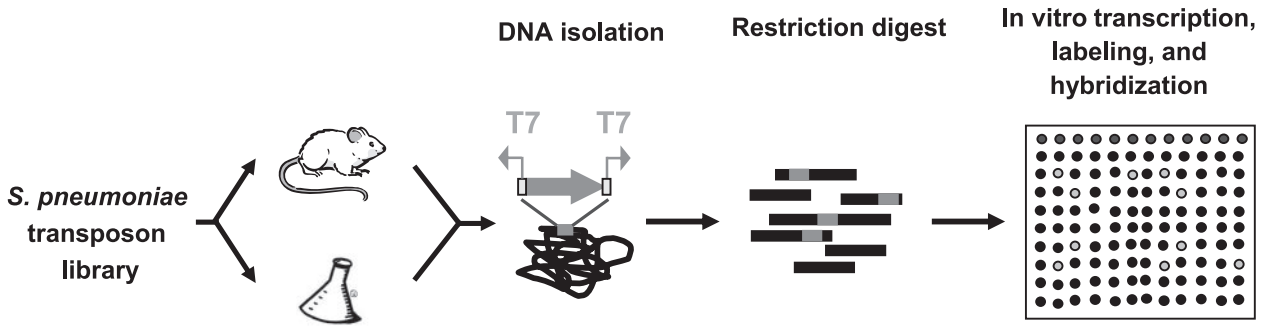

FIG. 1. Schematic representation of the GAF procedure. A large *S. pneumoniae* transposon library is grown under nonselective and selective conditions. Subsequently, chromosomal DNA containing transposons (gray rectangle) with outward-facing T7 RNA polymerase promoters (arrow with T7) is isolated from each population. The DNA is digested, and the DNA adjacent to the transposon insertion site is amplified using in vitro transcription with T7 RNA polymerase. The RNA is used in standard microarray procedures. Hybridization of these probes to a microarray will reveal which genes were disrupted in the mutants that disappeared during selection: those spots to which only material derived from the nonselective condition will hybridize (gray spots).

to the transposon insertion sites, as this determines the rate of false positives. Recently, several methods for identifying transposon insertion sites using microarrays have been described (6, 32, 33). In this paper, we describe (i) the pitfalls of several methods of amplification when applied to *S. pneumoniae* libraries and (ii) the successful development and validation of a specific and reproducible method for the detection of transposon insertion sites in *S. pneumoniae*. Finally we show, with a biological screen for genes essential for surviving zinc stress, that GAF is a reliable and effective method to identify conditionally essential genes.

MATERIALS AND METHODS

Bacterial strains, growth conditions, DNA isolation, and manipulations. The *S. pneumoniae* strains used in this study were the serotype type 2 isolate D39 (3) and its unencapsulated derivate R6 (8). *Lactococcus lactis* strain NZ9000 (10) was used to generate pGHT7:IS*S1*. *Escherichia coli* DH5αF was used as a host for the pR412 plasmid and the generation of pR412T7. All strains were routinely cultured in tryptone yeast broth or M17 broth containing glucose (GM17) as described previously (10). General DNA techniques were performed as described previously (10). Sequencing of the DNA adjacent to the insertion sites was performed using a single-primer PCR essentially as described previously (9). An in-frame deletion of *czcD* (SP1857) was constructed using pORI280 as described before (10) with primer pairs czcD-KO-1/czcD-KO-2 and czcD-KO-3/

TABLE 1. Primers used in this study[a]

Name	Sequence (5′–3′)
PBMrIRPi	AGACCGGGGACTTATCAGCC
PBMrStartT7	CCTATAGTGAGTCGTATTAGACACATAGATGGCGTCGC
PBMrEndT7	CCTATAGTGAGTCGTATTACAACATAGTTCCCTTCAAGA
ISSfor	ATTGTAAAACGACGGCCAGTGTTCATTGATATATCCTCGCTGTC
MATT1	AAGGGCATGGAAACAATTCGAGG
MATT2	GGAACGCTCTTCGGATTTTCGG
CEKG2A	CGCCACGCGTCGACTAGTACNNNNNNNNNNAGAG
CEKG2D	CGCCACGCGTCGACTAGTACNNNNNNNNNNATAG
CEKG2E	CGCCACGCGTCGACTAGTACNNNNNNNNNNGAAT
CEKG4	CGCCACGCGTCGACTAGTAC
LinkAmp1	AGGCACAGTCGAGGACTTATCCTA
LinkAmp2	GCCTCTGAATACTTTCAACAAGTTAC
PBMrTn5	CGGGAATCATTTGAAGGTTGGTACT
PBMrTn3	TACTAGCGACGCCATCTATGTGTC
TMr_1	TGCATTTAATACTAGCGACGCCATCTATGTGTC
IS*S1*_T7_up	CGATAAGCTTGATGGAGAGAATGGGTTCTGTTGCAAAGTTTTCTGATAAGTCTATTTTAGTGTAAAA TGAATAAATCTCCTATAGTGAGTCGTATTAAATGACAGCGAGGATATATCAATG
IS*S1*_T7_down	CAGGAATTCGATGCTCTAGAGCATTCTCTGGTTCTGTTGCAAAGTTTAAAAATCAAATACAAGGTCT CCTATAGTGAGTCGTATTATTTATAATCCTTCTTGTTCTTAAGCTAATATTCCC
Adapter1	AATTACCACGACCA
Adapter2	CGACCACGACCA
Adapter3	AGTCTCGCAGATGATAAGGTGGTCGTGGT
Adapter4	GTCCAGTCTCGCAGATGATAAGG
czcD-KO-1	TGCTCTAGAAGGTCAATGTCTCGATAAAG
czcD-KO-2	AGCATATTTTGCCTTCATATTTC
czcD-KO-3	ATATGAAGGCAAAATATGCTAGTTATGAGCATCAACATTAG
czcD-KO-4	GAAGATCTCTGTAGCTGAGACAAGCGC

[a] All primers were designed for this work, except CEKG2A and CEKG4, which were designed by N. R. Salama and coworkers (32).

czcD-KO-4 (Table 1), which removed 852 bp of the open reading frame (ORF); a detailed description of its construction is provided elsewhere (T. G. Kloosterman, M. M. Pol, J. J. E. Bijlsma, and O. P. Kuipers, submitted for publication).

Construction of pGh9:IS*S1* and pR412 derivates with outward-facing T7 RNA polymerase promoters in their transposons. A pGh9:IS*S1* derivative with two outward-facing T7 promoters in the IS*S1* element was constructed as follows. The IS*S1* element from pGh9:IS*S1* (21) was removed by digestion with EcoRI and HindIII and replaced by a HindIII-, EcoRI-digested PCR product of the IS*S1* element generated on pGh8:IS*S1* (21) with primers IS*S1*_T7_up (Table 1) and IS*S1*_T7_down (Table 1). This resulted in pGh9T7:IS*S1* which contains the IS*S1* element (with the T7 promoters on both sides) in an orientation opposite to that in the pGh9:IS*S1* plasmid.

The backbone of transposon donor plasmid pR412 (23), including the inverted left repeat and inverted right repeat of the transposon, was amplified with *Pwo* DNA polymerase (Roche) with a single phosphorylated primer PBMrIRPi (Table 1); PCR cycling conditions were as follows: 94°C for 4 min; 30 cycles of 94°C for 1 min, 55°C for 1 min, and 68°C for 5 min; and 68°C for 5 min. The spectinomycin-resistance cassette was PCR amplified with primers containing a T7 tag, i.e., PBMrStartT7 (Table 1) and primer PBMrEndT7 (Table 1); PCR cycling conditions were as follows: 94°C for 4 min; 30 cycles of 94°C for 1 min, 55°C for 1 min, and 68°C for 1.5 min; and 68°C for 4 min. The two PCR products were ligated to obtain the transposon donor plasmid pR412T7.

Generation of transposon libraries. Random transposon libraries were routinely generated in *S. pneumoniae* as described before, using the in vitro *mariner* transposon (1) and the in vivo pGh:IS*S1* system (21). Recombinant HimarC9 transposase was purified from the *E. coli* strain BL21(DE3) (Novagen) containing plasmid pET29C9 (13) essentially as described before (14). However, to prevent DNase contamination of the transposase stock, the DNase I treatment of the purification step was omitted from the protocol. For in vitro transposition, 1 μg of purified chromosomal DNA was incubated with 0.5 μg pR412T7, which contains the *marinerT7* transposon, and 0.5 μg purified recombinant HimarC9 transposase in a 20-μl reaction mixture. After repair of the transposition reaction with T4 DNA polymerase and *E. coli* DNA ligase (7), the DNA was used for transformation of strain R6 (1, 13). After overnight growth on selective plates containing 120 mg liter^{-1} spectinomycin, colonies were scraped, pooled, and stored at −80°C. For the construction of pGh9T7:IS*S1* mutant libraries (21), the R6 and D39 strains were transformed with the pGh9T7:IS*S1* plasmid and grown at 30°C on selective plates containing 0.25 mg liter^{-1} erythromycin. Several colonies of *S. pneumoniae* containing the pGh9T7:IS*S1* plasmid were pooled, grown overnight in GM17 broth with erythromycin at 30°C, and stored at −80°C. From this stock, bacteria were inoculated for overnight growth in GM17 broth at 30°C with the appropriate antibiotic, and the next day the cultures were diluted 1 to 50 in GM17 broth without antibiotics and after three hours of growth shifted to 37°C. After another three hours of growth at the nonpermissive temperature, 10^{-3} and 10^{-4} dilutions were plated on selective medium. Approximately 20,000 to 30,000 CFU were scraped from the plate, pooled, and stored at −80°C for further use. The same *marinerT7* and pGh9T7:IS*S1* transposon libraries were used for all experiments to determine the reproducibility and specificity of each method. New libraries were generated for the comparison of the *marinerT7* and pGh9T7:IS*S1* transposon libraries and the GAF zinc screen.

Microarray construction, labeling, hybridization, and analysis. *S. pneumoniae* microarrays were produced essentially as described previously (12, 40) and contained amplicons of on average 600 bp representing 2,087 ORFs of *S. pneumoniae* TIGR4, each of which are present twice on the array (36) as an internal control to monitor variations in hybridization efficiency per slide. Purified RNA obtained after in vitro transcription was used to generate fluorescent DNA probes by direct or indirect labeling using standard methods. Hybridization, washing, and acquisition of array images were performed as described previously (11). Spots were screened visually to identify those of low quality. The amount of low-quality spots varied per slide but was never more than 10%; they were removed from the data set prior to analysis. Slide data were processed and normalized using MicroPreP (41). A net mean intensity filter based on hybridization signals obtained with amplicons representing open reading frames unique for *S. pneumoniae* strain TIGR4 (36) was applied in all experiments. For one-array experiments (Fig. 2 and 3), only genes that had reliable measurements (i.e., passed filters) for both duplicate spots present on the array were analyzed. For the other experiments, microarray data were analyzed when at least three measurements were available, using a CyberT implementation of the Student's *t* test (20). False discovery rates were calculated as described previously (40). The correlation between normalized signal intensities per gene in both channels was determined by linear regression analysis.

Identification of transposon insertion sites using MATT. For the identification of transposon insertion sites, microarray tracking of transposon mutants (MATT) was performed as described previously (32) with some modifications. A linear PCR with AccuPrime (Invitrogen) polymerase was performed on chromosomal DNA containing pGh9T7:IS*S1* insertions as a template and the MATT1 (Table 1) primer located on the transposon, for 30 cycles at 56°C. The single-stranded fragments were purified with a GFX column (GE Healthcare), and 100 ng was used as a template for a semirandom PCR in a 20-μl reaction mixture containing 2 μM each of an anchored random primer (CEKG2A, CEKG2D, or CEKG2E; Table 1) and MATT1 and 1 U *Taq* polymerase (Invitrogen) with the following cycling reactions: 1 cycle at 94°C for 4 min; 12 cycles at 94°C for 30 s, 42°C for 30 s (−0.5°C each cycle), and 72°C for 2 min; and 30 cycles at 94°C for 30 s, 65°C for 30 s, and 72°C for 2 min. The reaction products were diluted fivefold, and 1 μl was used as a template for a subsequent reaction using a nested transposon-specific primer MATT2 (Table 1) and the anchor primer CEKG4 (Table 1) in a 100-μl reaction mixture using the following cycling conditions: 1 cycle at 94°C for 4 min, and 30 cycles at 94°C for 30 s, and 56°C for 30 s, and 72°C for 2 min. The PCR products were purified using a Roche PCR purification kit and used as input for a T7 RNA polymerase reaction using a T7 MEGAscript kit (Ambion).

Identification of transposon insertion sites using a TOPO adapter. Genomic DNA was diluted in 20 μl H_2O to 0.5 μg/μl and fragmented by sonication with the microtip of a Branson digital sonifier (eight times for 0.5 s, output 20%). To prepare double-stranded, blunt-end DNA, fragmented DNA was successively treated with S1 nuclease, T4 DNA polymerase, and *E. coli* DNA ligase. To facilitate the ligation of a TOPO-isomerase-activated linker (TOPO-Walker kit; Invitrogen), the blunt-end, double-stranded DNA was incubated with calf intestinal phosphatase for dephosphorylation and with *Taq* DNA polymerase (Thermoperfect; Integro) to create 3′ A-tailed DNA ends. Ligated DNA was used as a template in two consecutive PCRs with *Taq* DNA polymerase: the first reaction contained primers LinkAmp1 (Table 1) and PBMrTn5 (Table 1), and the second contained nested PCR primers LinkAmp2 (Table 1) and PBMrTn3 (Table 1). PCR cycling conditions for both reactions were as follows: 94°C for 4 min; 20 cycles of 94°C for 30 s, 63°C for 1 min, and 72°C for 1.5 min; and 72°C for 3 min. PCR products were loaded on a 2% agarose gel, and DNA fragments with the preferred size were excised and extracted with a Qiaquick gel purification kit (QIAGEN). In vitro transcription reaction and subsequent DNase I digestion were performed with a T7 MEGAscript kit (Ambion).

Identification of transposon insertion sites using TraSH. The TraSH method was performed essentially as described previously (33) with a few modifications. Genomic DNA was digested with ApoI or TacI and precipitated and ligated overnight in a 15-μl volume with an approximately 1,000-fold molar excess of the appropriate adapter (Adapter1 for ApoI- and Adapter2 for TacI-digested DNA; Table 1). The ligated products were purified, and 0.3 to 1 μl was used as a template in the following PCR containing primer TmR1 (Table 1): 1 cycle at 95°C for 1 min; and 5 cycles at 95°C for 30 s, 72°C for 30 s, and 72°C for 1 min. Subsequently the appropriate adapter primer (Adapter3 for ApoI- and Adapter4 for TacI-digested DNA; Table 1) was added, and the reaction was continued with the following cycling conditions: 5 cycles at 95°C for 30 s, 72°C for 30 s, and 72°C for 1 min, and 30 cycles at 94°C for 30 s, 65°C for 30 s (−0.3°C each cycle), and 72°C for 70 s. PCR products were purified and used as a template for a T7 RNA polymerase reaction using a MEGAscript kit (Ambion).

Identification of transposon insertion sites using the T7 RNA polymerase promoter. The procedure was performed essentially as described previously (6). Genomic DNA (>10 μg) was digested overnight with TacI, DdeI, or AluI, and the purified DNA was used for an in vitro transcription reaction with a MEGAscript T7 RNA polymerase kit (Ambion) in a total volume of 20 μl.

Selection for genes essential for surviving high-zinc conditions. Aliquots of a *marinerT7* transposon library generated in *S. pneumoniae* R6 containing approximately 20,000 independent mutants were diluted 1 to 20 in GM17 broth or in GM17 broth containing 0.5 mM $ZnSO_4$. Both cultures were grown for 10 generations as determined by the optical density at 600 nm, after which the bacteria were spun down (9,000 × *g*, 4 min) and used for the extraction of genomic DNA.

Microarray data accession numbers. The microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under GEO Series accession numbers GSM158348, GSM158350, GSM158558, GSM158562, GSM158563, GSM158564, GSM158565, GSM158566, GSM158567, GSM158568, GSM158569, and GSM158570.

RESULTS

Prerequisites for the development of GAF. Important considerations for the development of GAF were that the method

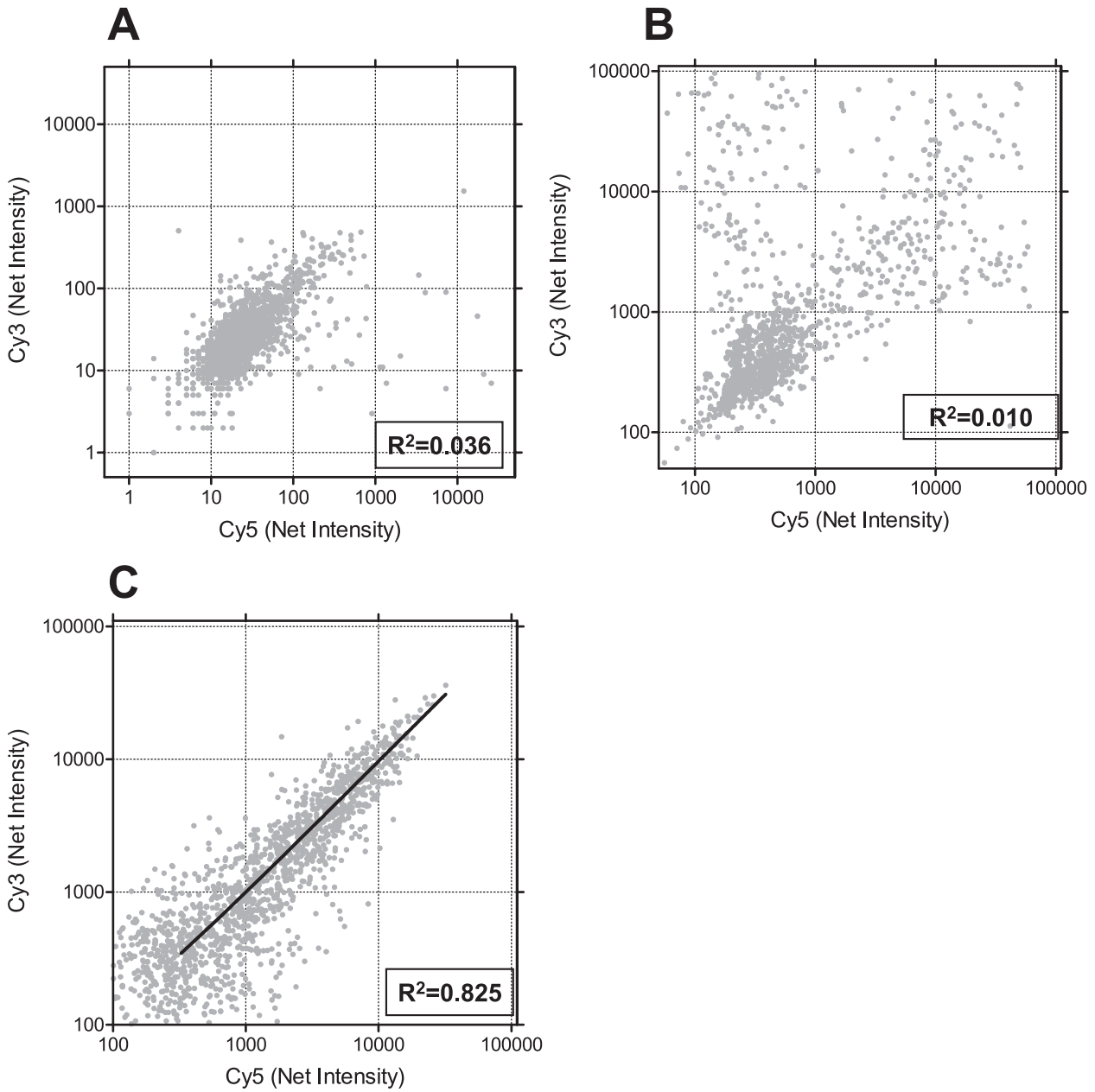

FIG. 2. Detection of transposon insertion sites using PCR-based methods. (A) Scatterplot of the signal intensities of two parallel reactions on the chromosomal DNA of a genome-wide library using MATT. There is no correlation between the signal intensities for each gene (r^2, 0.036), indicating that the procedure is nonreproducible. (B) Scatterplot of the signal intensities of two parallel reactions on the chromosomal DNA of a genome-wide library using a TOPO-isomerase adapter combined with a nested PCR. There is no correlation between the signal intensities in each gene (r^2, 0.010), indicating that the procedure is nonreproducible. (C) Scatterplot of the signal intensities of two parallel reactions on the chromosomal DNA of a genome-wide library using TraSH. Correlation between the signal intensities for each gene indicates that the procedure is fairly reproducible (r^2, 0.825).

should be sensitive enough to detect all mutants in a library and should be reproducible. Reproducibility, specificity, and sensitivity are in large part determined by the amplification and detection step (Fig. 1). Therefore, we tested the performance of various methods for transposon insertion site detection (6, 32, 33) on both small (<100 CFU) and large (>20,000 CFU) mutant libraries.

Although we used strains R6 and D39 throughout this study, we wanted our method to be applicable to (in principle) all *S. pneumoniae* strains, some of which are not as easily transformable. Therefore, we established two methods for the generation of random mutant libraries: in vitro *mariner* mutagenesis (1), which depends on a high-transformation efficiency of the accepting strain for the generation of large libraries and is not expected to have polar effects on downstream genes (31), and the in vivo pGh:IS*S1* system (21), which consists of a plasmid with a temperature-sensitive origin of replication, the IS*S1* insertion element, and an erythromycin resistance gene with a

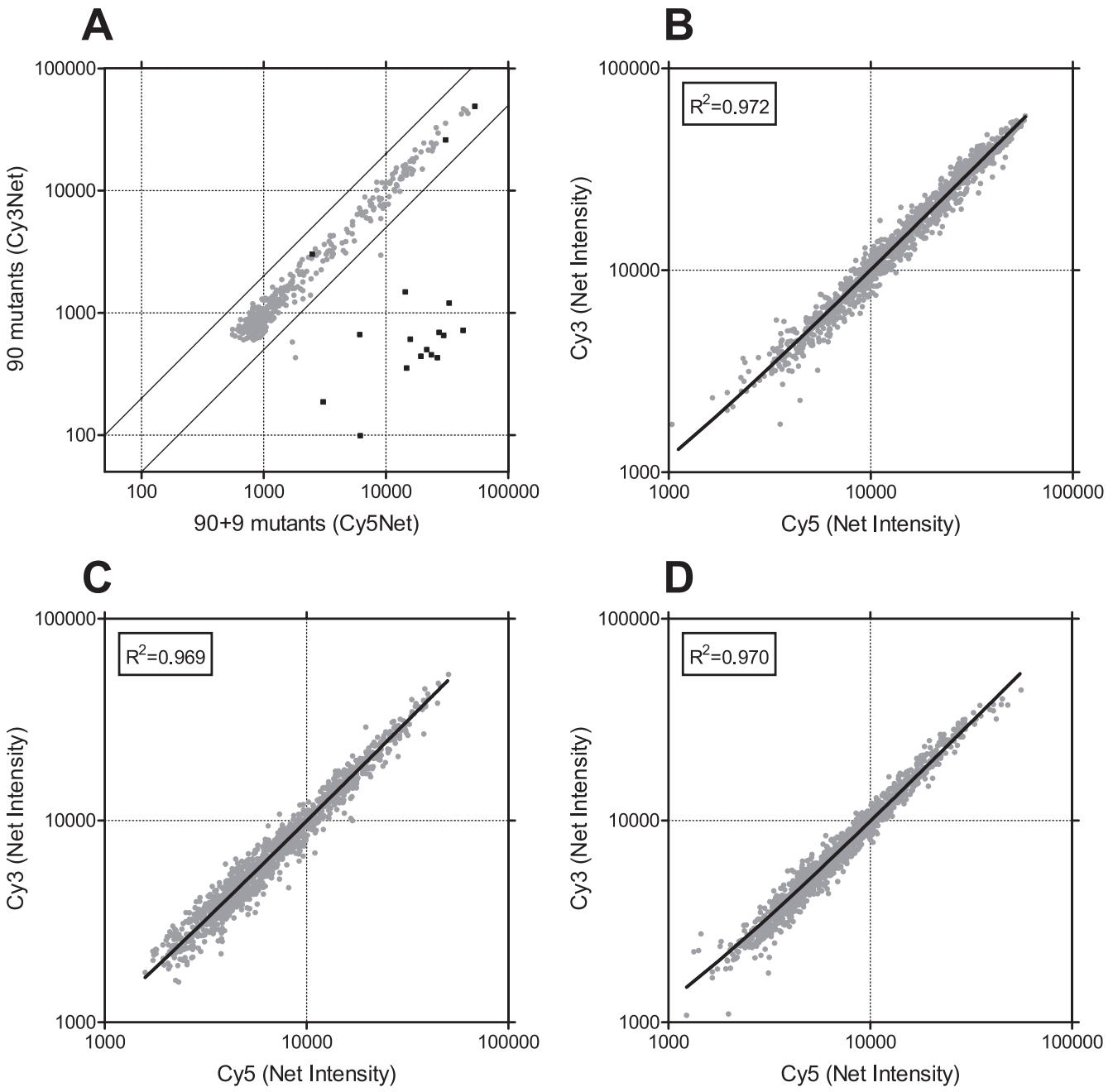

FIG. 3. In vitro transcription on purified chromosomal DNA restriction fragments leads to specific and sensitive detection of transposon insertion sites. (A) Scatterplot of the comparison of reactions on the chromosomal DNA of a mutant library consisting of 90 unique *marinerT7* mutants and the DNA of the same mutants to which the chromosomal DNA of 9 defined mutants has been added. Straight lines denote a twofold difference in signal intensities between the Cy5 and Cy3 channels. Genes identified by sequencing as added to the library are indicated as black squares. (B to D) Scatterplots of the signal intensities of two parallel reactions on the chromosomal DNA of a genome-wide library digested with TacI (B), DdeI (C), and AluI (D). Correlation between the signal intensities in each plot is >0.96, indicating that the procedure is highly reproducible.

weak terminator (-7.5 kcal mol^{-1}); thus, insertion of pGh:IS*S1* could cause some polar effects. Growth of bacteria that contain the plasmid for several generations at 30°C, the permissive temperature, followed by a temperature shift to 37°C, induces the insertion of the plasmid into the chromosome. Thus, a one-time introduction of this plasmid in a strain with low transformation efficiency is sufficient to generate large (>20,000 CFU) libraries, avoiding the need for high-transformation efficiencies in accepting strains. In each transposon, an outward-facing T7 RNA polymerase promoter close to both ends of the insertion element was introduced for amplification purposes. Both transposons readily generated large (>20,000 CFU) libraries in *S. pneumoniae* strains R6 and D39, as expected, and in strain TIGR4 with the pGh:IS*S1* system (34) as

well (results not shown). Mutant libraries generated with these transposons were used exclusively throughout this study, and the RNA generated with the aid of the T7 RNA polymerase promoter was used in our standard microarray procedures.

Identification of the transposon insertion sites using MATT. Microarray tracking of transposon mutants combines two rounds of PCR: first, a reaction with specific and nonspecific primers, followed by a nested PCR to further amplify the material (32). The specific primers hybridize to sequences on the transposon, and the nonspecific primers contain 10 degenerative nucleotides followed by a unique anchor sequence. Several anchor sequences were chosen that are present on average every 400 bp in the genome. MATT was found to be specific and reproducible when used for artificial GAF experiments on a few mutants (comparison of 12 mutants with the same set in which three mutants were removed; results not shown). However, when two parallel reactions on one large (>20,000 CFU) genome-wide library were compared, it was obvious that the method was not reproducible (Fig. 2A), as few genes displayed equal signal intensities in both channels (r^2, 0.036). When the products of the second round of PCR were visualized on an agarose gel, several distinct bands of various sizes in each reaction were visible, indicating that some DNA fragments were selected randomly during the PCR steps (results not shown).

Identification of transposon insertion sites using a TOPO-isomerase adapter or the adapter-based TraSH method. Next, we tested a method that consists of the ligation of a TOPO-isomerase-activated adapter to sheared chromosomal DNA, followed by two PCR steps using nested primers that anneal specifically to the adapter and the transposon. Similar to MATT, this method was specific and reproducible when applied to artificial GAF experiments on a small number of mutants (i.e., comparison of 21 mutants with the same set from which 6 mutants were removed; results not shown) but was clearly nonreproducible when used on a genome-wide library (Fig. 2B). When the PCR products of separate reactions were examined, on an agarose gel, different distinct banding patterns were visible, again indicating that the PCR steps introduced a bias in the whole procedure. Repeated attempts to improve the PCR steps in both the MATT and the TOPO-adapter method by using *Taq* polymerase variants with increased specificity, touchdown PCR, and even more stringent annealing conditions were unsuccessful (results not shown).

Subsequently, we tested the TraSH method, which consists of the ligation of an adapter to chromosomal DNA digested with TaqI or ApoI, followed by a short linear PCR with a transposon-specific primer and a PCR with transposon-specific and adapter-specific primers (33). This method was fairly specific and reproducible both when used on a small number of mutants (i.e., comparison of 21 mutants with the same set of mutants from which 6 mutants were removed; results not shown) and on a large (>20,000 CFU) genome-wide library (Fig. 2C), as the majority of the genes had comparable signal intensities in both channels on two parallel reactions (r^2, 0.83). However, the optimal annealing temperature and template concentration had to be determined for each new sample, as standardized PCR conditions resulted in distinct banding patterns in the PCR products and lower correlations between signal intensities in each channel. As the continuous optimization of the PCR conditions made this method cumbersome, we did not pursue it further.

Identification of transposon insertion sites using in vitro transcription on digested DNA. To avoid the bias introduced by methods employing PCR, we directly used purified chromosomal fragments digested with TaqI (which on average cuts every 442 bp) as input for an in vitro transcription reaction using the T7 RNA polymerase promoter that we had introduced in both transposons (Fig. 1) (6, 17). Comparison of duplicate reactions on a small set of 90 mutants with this method gave rise to specific and reproducible results (data not shown). Next, we tested the feasibility of this method by performing an artificial GAF experiment. To a small library of 90 mutants, 9 defined mutants (transposon insertion site determined by sequencing) were added (90 + 9 library). Subsequently, transposon insertion sites in both libraries were determined by microarray analysis as described above. As expected, the signal intensities of most genes were found to be similar in both channels (i.e., ratios between 0.5 and 2), indicating that mutants in these genes were present in both populations. Furthermore, the signal intensities of 13 genes in the 90 + 9 library were increased significantly in one channel (Fig. 3A), leading to ratios of >9 (Table 2), indicating that transposon insertions in or near these genes were not present in the 90 mutant library. All 13 of these genes corresponded to ORFs present on the TaqI fragments of 8 defined mutants added to the library (Fig. 3A; Table 2). The corresponding amplicons of the genes on the TaqI restriction fragment in one defined mutant (SP2129 to SP2131) had high signal intensities in both channels (ratio of 1), indicating that a mutant with a transposon insertion in this location was already present in the 90 mutant library. Three additional genes were detected as having been added to the library (Fig. 3A; Table 2). The detection of SP1792 is probably due to the fact that the amplicon has significant homology with, among others, ORF SP1262 (data not shown), and thus the high Cy5 signal is probably due to cross hybridization with the fragments generated by the added mutant A2. SP0001 is adjacent to ORF SP2240 on the genome and is present on the TaqI fragment of added mutant A10, and the intermediate-level signals in the Cy5 channel could be due to the occurrence of some partially digested fragments. It is not clear why gene SP2169 is detected; however, the signals in both channels and the ratio are close to the arbitrarily chosen cutoff value of 2 and differ substantially from the signal intensities of the amplicons that represent the 9 added mutants.

Subsequently, this procedure was tested on one genome-wide *marinerT7* library (>100,000 CFU) grown for approximately 20 generations in GM17. Three different restriction enzymes were used, TaqI, DdeI, and AluI, which have an increasing number of sites in the *S. pneumoniae* genome; on average, they cut every 442, 337, and 204 bp, respectively. The scatterplots derived from the analysis of each separate hybridization (Fig. 3B, C, and D) clearly show a strong correlation between the signal intensities in each channel for each gene (r^2 of >0.96 in all cases) (the normalized signal data and CyberT output data for these arrays are available at http://molgen.biol.rug.nl/publication/GAF_data/). Restriction with DdeI resulted in the detection of 1,301 genes, and the numbers for the TaqI and AluI digests were comparable (1,193 and 1,265, re-

TABLE 2. ORFs present on the TaqI fragments that contain the insertion sites of the 9 mutants added to the library and the corresponding normalized signals and ratios[a]

Mutant	ORFs present on TaqI fragment	Gene ID on the array	Cy5 signal	Cy3 signal	Ratio (90 + 9 mutants/90 mutants)
A1	SP0319	SP0319	15761	600	25.9
	SP0320	SP0320	14689	354	41.5
A2	SP1262	NP[b]			
	SP0029	SP0029	6076	665	9.1
	SP0031	NP			
	SP0032	SP0032	14318	1484	9.6
A4	SP0943	SP0943	19237	441	43.6
A6	SP2001	SP2001	6152	99	62.1
	SP2002	SP2002	27091	692	39.1
	SP2003	ND	ND	ND	
A8	SP1533	SP1533	32682	1204	27.2
A10	SP2239	SP2239	ND	ND	ND
	SP2240	SP2240	21523	501	42.9
B7	SP2151	SP2151	3063	187	16.4
	SP2152	SP2152	42633	718	59.4
	SP2153	SP2153	23473	454	51.7
B8	SP2129	SP2129	53211	48992	1.1
	SP2130	SP2130	2486	3021	0.8
	SP2131	SP2131	30685	25927	1.2
C2	SP1999	SP1999	26273	430	61.0
NA	NA	SP1792	9058	2955	3.1
NA	NA	SP2169	1715	577	3.0
NA	NA	SP0001	1814	430	4.2

[a] ND, not detected (signal below background filter); NA, not applicable.
[b] NP, ORF not represented by an amplicon on the microarray.

spectively). Therefore, the three duplicate hybridizations were combined for further analysis. The library was generated in R6, which contains 2,116 ORFs, 1,995 of which are homologous to ORFs in strain TIGR4, and thus are present on our array (8). As expected, all amplicons that represent ORFs present in strain TIGR4 but not in R6 (36) (TIGR4 specific) had low signal intensities. These signals were used to manually generate a background filter that was applied to the analysis of the data from all three restrictions. Of the 1,995 R6 ORFs on the array, 1,322 (66%) were detected in at least two restrictions and had a less than twofold ratio difference, demonstrating the reproducibility of this method. In principle, no mutants in genes that are essential for the growth of *S. pneumoniae* can be present in the library; thus, amplicons representing these genes should not give signals above background (i.e., should be classified as nondetected). Comparison of the 673 nondetected R6 genes with the reported essential genes of *S. pneumoniae* (1, 5, 15, 18, 25, 35, 37, 43) showed that 80 had been designated as essential and 97 were located adjacent to putative essential genes (see Table S1 in the supplemental material). These analyses showed that this method was robust and reproducible for incorporation in GAF and that all three restriction enzymes can be used. The latter is important, as not all restriction enzymes cut in all genes and the use of several restriction enzymes on the same DNA improves the chance of detecting all mutants that have disappeared from a population.

Comparison of the pGhT7:IS*S1* and *marinerT7* transposon libraries. Previous experiments were performed on *marinerT7*-generated libraries. To test whether the method could also be applied to pGh9T7:IS*S1* libraries, and to assess the level of saturation and randomness achieved by each transposon, we compared a large (>20,000 CFU) *marinerT7* library with a pGh9T7:IS*S1* one (>20,000 CFU), each grown for approximately 20 generations in GM17. After isolation, the DNA was digested with DdeI and AluI and each detection reaction was performed in duplicate. Microarray data were analyzed as before by using the background filter generated with the TIGR4-specific amplicons (the normalized signal data and CyberT output data for these arrays are available at http://molgen.biol.rug.nl/publication/GAF_data). This analysis resulted in the detection of 1,473 genes, and the nondetected genes were similar to those not detected in the analysis of the three *marinerT7* digests. The ratio of the majority of the genes was between 0.5 and 2, indicating that they are mutated by both transposons. However, 139 (9%) genes had a ratio lower than 0.5 or higher than 2, indicating that these genes are only, or preferentially, mutagenized by one of the two transposons. These "hot spots" are distributed throughout the genome (Table 3). There were 60 genes preferentially mutated by pGh9T7:IS*S1*, whereas 79 were preferentially mutated by *marinerT7*, indicating that it has a slightly higher saturation rate than pGh9T7:IS*S1*. Anchored PCR using primers located on some of the ORFs that had a ratio close to the cutoff point of 0.5 indicated that these genes were indeed preferentially hit by either transposon and that these ratios were not due to slight differences in, for instance, T7 RNA polymerase efficiency (data not shown).

Identification of genes essential for surviving zinc stress. As all components for GAF were successfully developed, we tested its ability to identify conditionally essential genes from a library grown under in vitro stress conditions. Although Zn^{2+} is an essential ion, it is toxic to bacteria in high concentrations. Bacteria often contain specific proteins for Zn^{2+} secretion (26), but none have been identified in *S. pneumoniae* so far.

TABLE 3. Summary of genes preferentially mutated by *marinerT7* (ratio of <0.5 for the first two columns) or by pGh9T7:IS*S1* (ratio of >2 for the last two columns)[a]

Gene ID	Ratio (pGh9T7:IS*S1*/*marinerT7*)	Gene ID	Ratio (pGh9T7:IS*S1*/*marinerT7*)
SP0024–SP0027	0.40; 0.50; 0.26; 0.32	SP0054	2.45
SP0115	0.48	SP0069	2.37
SP0151	0.47	SP0071	2.52
SP0240	0.45	SP0097	2.15
SP0242	0.13	SP0105	2.03
SP0278	0.42	SP0112-SP0113	3.05; 3.52
SP0319	0.44	SP0116	2.62
SP0321	0.45	SP0123	2.15
SP0322	0.46	SP0340	2.03
SP0385	0.41	SP0412	2.06
SP0558	0.49	SP0439	2.41
SP0574	0.47	SP0530	2.19
SP0593	0.47	SP0585	2.50
SP0690–SP0693	0.34; 0.33; 0.33; 0.38	SP0609	2.43
SP0741–SP0743	0.13; 0.18; 0.35	SP0617-SP0618	2.65; 2.03
SP0846-SP0847	0.31; 0.29	SP0681	2.20
SP0943	0.28	SP0764-SP765	2.92; 2.11
SP0987-SP0988	0.41; 0.49	SP0800	2.50
SP1025	0.44	SP0881-SP0882	2.99; 2.61
SP1027	0.37	SP1018	2.08
SP1035	0.44	SP1104-SP1105	2.22; 2.04
SP1040–SP1043	0.35; 0.27; 0.47; 0.35	SP1246	2.16
SP1045	0.48	SP1249	2.23
SP1051	0.47	SP1376	2.26
SP1053	0.47	SP1505	2.83
SP1069	0.45	SP1578	2.04
SP1089	0.35	SP1586	2.16
SP1116	0.48	SP1609	2.70
SP1121	0.47	SP1627	2.07
SP1122	0.46	SP1659	3.00
SP1159–SP1164	0.47; 0.11; 0.38; 0.33; 0.34; 0.35	SP1691	2.11
SP1166	0.45	SP1743	2.13
SP1170-SP1171	0.49; 0.35	SP1790	2.56
SP1173	0.39	SP1795	4.56
SP1184	0.49	SP1852	2.74
SP1186-SP1187	0.49; 0.26	SP1981–SP1983	2.95; 5.43; 3.45
SP1190–SP1192	0.33; 0.40; 0.50	SP1996	2.04
SP1215	0.46	SP2010	2.39
SP1222	0.35	SP2031-SP2032	2.31; 2.14
SP1258	0.45	SP2070–SP2072	3.14; 2.40; 4.22
SP1275	0.47	SP2077	2.48
SP1436	0.47	SP2082	2.55
SP1552	0.46	SP2156	2.77
SP1640	0.48	SP2170	2.08
SP1693	0.35	SP2192-SP2193	2.30; 2.23
SP1815	0.48	SP2205-SP2206	2.85; 2.77
SP1856	0.49	SP2235-SP2236	2.04; 2.17
SP1897-SP1898	0.33; 0.49		
SP1923-SP1924	0.27; 0.46		
SP1954–SP1957	0.25; 0.28; 0.34; 0.40		
SP1959	0.42		
SP2098	0.28		

[a] Sorted by their location on the genome to show clustering. Gene ID refers to TIGR4 locus tags. Bayesian *P* value of <0.0001 by CyberT, false discovery rate of <0.002 for all genes.

The concentration of Zn^{2+} in serum is reported to be around 15.3 μM (1 mg liter^{-1}) and in lung tissue 229.4 μM (15 μg/g wet tissue) (42), and during inflammation concentrations increase in blood and other body sites (24, 38), indicating that *S. pneumoniae* is likely to encounter Zn^{2+} stress during infection. Therefore, we decided to screen for genes essential for the survival of Zn^{2+} stress. In addition, a zinc exporter had already been identified by other means by our group (T. G. Kloosterman et al., submitted). A large (>20,000 CFU) *marinerT7* library in strain R6 was grown in quadruplicate in GM17 or GM17 containing 500 μM $ZnSO_4$, which is half the MIC of zinc for *S. pneumoniae* (data not shown), for approximately 10 generations. Subsequent detection of transposon insertion sites in both populations using a TaqI restriction showed that gene SP1857 and a cluster of two genes, SP0856-SP0857, had a significantly lower ratio (Table 4), suggesting that these are essential genes for growth in high-Zn^{2+} concentrations. SP0856-SP0857 are annotated as *ilvE* and as an oligopeptide-binding

TABLE 4. Summary of genes essential for surviving high-Zn^{2+} concentrations identified with GAF[a]

Gene ID	Function	Ratio (GM17 with Zn^{2+}/GM17)
SP1857	Cation efflux system protein CzcD	0.43
SP0856	Branched-chain amino acid aminotransferase IlvE	0.47
SP0857	ABC transporter substrate binding protein	0.48

[a] Genes were designated conditionally essential when the GM17/GM17 with Zn^{2+} ratio was <0.5, the Bayesian *P* value was <0.001, and the false discovery rate was <0.05. Gene ID refers to TIGR4 locus tags; function is based on the TIGR annotation (36).

protein, respectively. SP1857, annotated as *czcD*, encodes a protein homologous to cation exporters (36) and was the exporter already identified as being involved in Zn^{2+} resistance (T. G. Kloosterman et al., submitted). A mutant generated in this gene was indeed unable to grow in GM17 containing 500 μM $ZnSO_4$ (Fig. 4A and B). Although the mutant was at least ten times more sensitive to $ZnSO_4$ than the parent strain, the ratio in the GAF experiment was 0.43. In silico TacI restriction analysis of the R6 nucleotide sequence of the homologue of SP1857 and the chromosomal region surrounding it results in a DNA fragment that contains SP1857, its promoter, and part of the adjacent gene, SP1858 (Fig. 4C). As polar effects are unlikely to occur, mutants containing a transposon inserted in front of the *czcD* ORF probably have no growth defect in GM17 plus $ZnSO_4$ and will not disappear from the population. The presence of such mutants in the library will result in a DNA fragment that also hybridizes to the SP1857 amplicon, which can explain why the ratio was close to 0.5. However, this ratio was found to be significant by statistical analysis, and the phenotype of the deletion mutant confirmed our GAF findings. Thus, we have successfully developed GAF, as it correctly identifies conditionally essential genes in *S. pneumoniae*.

DISCUSSION

Here, we describe the development of GAF for *S. pneumoniae*, a technique enabling genome-wide negative screens in a high-throughput fashion. Selection for genes essential to survive high-Zn^{2+} concentrations with GAF correctly identified *czcD* as the main determinant for this resistance, which was also shown in a study exploring the response of *S. pneumoniae* to zinc stress (T. G. Kloosterman et al., submitted). One of the other genes identified in this screen, SP0857, encodes a putative ABC transporter, suggesting that it could be involved in the secretion of Zn^{2+}. The function of *ilvE* (SP0856) in resistance to Zn^{2+} stress is not immediately obvious; it could either be that mutation of this gene confers a growth defect under these particular conditions or that there is a polar effect of the insertion on SP0856 or the other two downstream genes that encode membrane proteins. The function of these genes in Zn^{2+} stress is currently under investigation. GAF has also been applied successfully to identify genes essential for natural competence (our unpublished results), and we are currently using it to identify genes essential in several different disease aspects.

GAF has several advantages. Detection of transposon insertion sites in a library is performed in a genome-wide manner using microarrays, bypassing the need to generate mutants with unique tags and/or store all mutants separately; this allows for the use of large mutant libraries that have more than 10× coverage of the genome. It also eliminates the need to sequence each individual mutant to determine the transposon insertion site, which accelerates the whole screening procedure and makes it possible to go from recovering the mutants after selection to identifying all transposon insertion sites in about one week.

Recently, several methods have been developed to identify transposon insertion sites using microarrays (32, 33, 39). For all these approaches it is important to reduce the number of false positives, as the mutants are not stored separately and site-directed mutants have to be generated to confirm that the identified genes are indeed conditionally essential. Therefore, we have looked carefully into the specificity and reproducibility

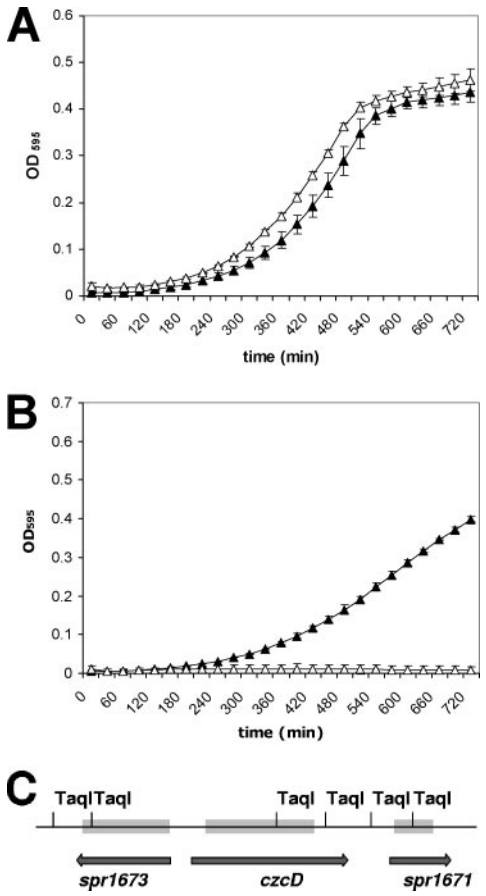

FIG. 4. *czcD* is a conditionally essential gene when *S. pneumoniae* is grown under high-Zn^{2+} conditions. (A) Growth of the wild type (filled triangles) and a *czcD* deletion mutant (open triangles) in GM17. (B) Growth of the wild type (filled triangles) and a *czcD* deletion mutant (open triangles) in GM17 to which 0.5 mM $ZnSO_4$ was added. Results are representative of at least three replicate experiments; the error bars indicate the standard deviations. OD_{595}, optical density at 595 nm. (C) Schematic representation of in silico digestion of the R6 chromosomal region containing *czcD* and its neighboring genes with TaqI. Arrows indicate the locations of genes, and gray rectangles indicate the genomic regions that are present as amplicons on the microarray.

of these methods, an issue that is rarely addressed. We showed that methods that use a PCR step or a combination of PCR and adapters introduce a random selection of certain fragments in the procedure when used on genome-wide libraries, which leads to nonreproducible results. A similar effect was observed by Tong et al., and they showed that this is not dependent on the use of amplicon arrays (39). It seems more likely that this random selection of certain fragments is caused by the presence of more than 500 mutants at once, as sensitive and specific results with PCR-based methods can be obtained when small libraries (<500 CFU) or subpools of a large library are used (39; results not shown).

When the PCR step is omitted and restricted DNA is used directly as a template for an in vitro transcription reaction, the detection of transposon insertion sites becomes reproducible and specific (Fig. 3). We used the signal present on TIGR4-specific amplicons for the generation of a background filter. Application of this filter resulted in the detection of 66% of the ORFs of strain R6 on our array, which is in agreement with the proposed *S. pneumoniae* core genome consisting of 73% of the TIGR4 ORFs (28). As expected, the set of nondetected genes contained many essential genes or genes adjacent to essential genes. The latter indicates that there are quite a number of polar effects in the studies investigating the essential genes of *S. pneumoniae*. Other nondetected genes are either hypothetical or expected to be essential as they, for instance, encode parts of the ribosome (e.g., SP0213, SP0214, and SP0232-SP0233-SP0234; see Table S1 in the supplemental material). A few undetected genes are represented by short amplicons, and it could be that insertions in this region, combined with the restriction enzymes used, do not generate fragments that hybridize to these amplicons (this could be the case for, e.g., SP0540 to SP0548, which encode a putative bacteriocin operon, and we do not expect these genes to be essential; see Table S1 in the supplemental material). We also showed that large genome-wide libraries (>20,000 CFU) generated with *marinerT7* and pGh9T7:IS*S1* have a similar degree of saturation but that each transposon is missing approximately 4% of the total number of genes. These "blind spots" are probably due to differences in the insertion site preferences of each transposon, although the genes missed by *marinerT7* do contain its recognition site TA. This also shows that the duplication of pGh9T7:IS*S1* in the genome (21) does not interfere with the detection of the transposon insertion sites and that this system is a good alternative for the *marinerT7* system for use in strains that do not have a high-transformation efficiency.

In contrast to transcriptome analysis, the severalfold change obtained by GAF analysis bears no direct relation to the functional impairment caused by mutation of a gene. For instance, the ratio of *czcD* in the Zn^{2+} stress GAF screen was 0.43, but a deletion mutant was at least ten times more sensitive to high-Zn^{2+} concentrations than the wild type. This apparent discrepancy has two probable causes. (i) The ratio is dependent on the generated restriction fragments, which prevents control over the length of the probes generated for hybridization. It is conceivable that an essential gene is also represented on a fragment derived from a transposon insertion in an adjacent, nonessential gene, thereby masking the effect of the disappearance of the mutant in the conditionally essential gene. This is probably the case for *czcD*. (ii) Not all transposon insertions abolish gene function completely; however, these mutants will also generate a signal during the GAF procedure, again leading to a masking of the effect of the complete disruption of gene function. These effects are amplified when large, complex libraries are used, which explains why the ratios in the artificial go + g library GAF experiment are much higher than those obtained when genome-wide libraries are used.

In conclusion, GAF is a versatile, high-throughput method for the identification of conditionally essential genes in *S. pneumoniae*. GAF should be easily adaptable to other microorganisms, as the only prerequisites for this technique are the availability of random (transposon) mutant libraries and a microarray.

ACKNOWLEDGMENTS

We thank Marc Prudhomme and Jean-Pierre Claverys (CNRS, Toulouse, France) and Andrew Camilli (Tufts University School of Medicine, Boston, MA) for kindly providing plasmids pR412 and pET29C9. We thank Don Morrison (University of Chicago) for his generous gift of competence-stimulating peptide. We thank Magda Pol for skillful technical assistance with the Zn^{2+} growth experiments.

This work was supported by IOP genomic grant IGE3002 of the Dutch Ministry of Economic Affairs.

REFERENCES

1. **Akerley, B. J., E. J. Rubin, A. Camilli, D. J. Lampe, H. M. Robertson, and J. J. Mekalanos.** 1998. Systematic identification of essential genes by in vitro *mariner* mutagenesis. Proc. Natl. Acad. Sci. USA **95:**8927–8932.
2. **Autret, N., and A. Charbit.** 2005. Lessons from signature-tagged mutagenesis on the infectious mechanisms of pathogenic bacteria. FEMS Microbiol. Rev. **29:**703–717.
3. **Avery, O. T., and R. Dubos.** 1931. The protective action of a specific enzyme against type III pneumococcus infection in mice. J. Exp. Med. **54:**73–89.
4. **Bogaert, D., R. de Groot, and P. W. Hermans.** 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. Lancet Infect. Dis. **4:**144–154.
5. **Carver, T. E., B. Bordeau, M. D. Cummings, E. C. Petrella, M. J. Pucci, L. E. Zawadzke, B. A. Dougherty, J. A. Tredup, J. W. Bryson, J. Yanchunas, Jr., M. L. Doyle, M. R. Witmer, M. I. Nelen, R. L. DesJarlais, E. P. Jaeger, H. Devine, E. D. Asel, B. A. Springer, R. Bone, F. R. Salemme, and M. J. Todd.** 2005. Decrypting the biochemical function of an essential gene from *Streptococcus pneumoniae* using ThermoFluor technology. J. Biol. Chem. **280:**11704–11712.
6. **Chan, K., C. C. Kim, and S. Falkow.** 2005. Microarray-based detection of *Salmonella enterica* serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. Infect. Immun. **73:**5438–5449.
7. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389–1406.
8. **Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. Leblanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass.** 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709–5717.
9. **Karlyshev, A. V., M. J. Pallen, and B. W. Wren.** 2000. Single-primer PCR procedure for rapid identification of transposon insertion sites. BioTechniques **28:**1078–1082.
10. **Kloosterman, T. G., J. J. Bijlsma, J. Kok, and O. P. Kuipers.** 2006. To have neighbour's fare: extending the molecular toolbox for *Streptococcus pneumoniae*. Microbiology **152:**351–359.
11. **Kloosterman, T. G., W. T. Hendriksen, J. J. Bijlsma, H. J. Bootsma, S. A. van Hijum, J. Kok, P. W. Hermans, and O. P. Kuipers.** 2006. Regulation of glutamine and glutamate metabolism by GlnR and GlnA in *Streptococcus pneumoniae*. J. Biol. Chem. **281:**25097.
12. **Kuipers, O. P., A. de Jong, R. J. Baerends, S. A. van Hijum, A. L. Zomer, H. A. Karsens, C. D. den Hengst, N. E. Kramer, G. Buist, and J. Kok.** 2002. Transcriptome analysis and related databases of *Lactococcus lactis*. Antonie Leeuwenhoek **82:**113–122.
13. **Lampe, D. J., B. J. Akerley, E. J. Rubin, J. J. Mekalanos, and H. M. Robertson.** 1999. Hyperactive transposase mutants of the Himar1 *mariner* transposon. Proc. Natl. Acad. Sci. USA **96:**11428–11433.

14. **Lampe, D. J., M. E. Churchill, and H. M. Robertson.** 1996. A purified *mariner* transposase is sufficient to mediate transposition in vitro. EMBO J. **15:**5470–5479.
15. **Lara, B., A. I. Rico, S. Petruzzelli, A. Santona, J. Dumas, J. Biton, M. Vicente, J. Mingorance, and O. Massidda.** 2005. Cell division in cocci: localization and properties of the *Streptococcus pneumoniae* FtsA protein. Mol. Microbiol. **55:**699–711.
16. **Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison, and D. W. Holden.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:** 555–571.
17. **Lawley, T. D., K. Chan, L. J. Thompson, C. C. Kim, G. R. Govoni, and D. M. Monack.** 2006. Genome-wide screen for salmonella genes required for long-term systemic infection of the mouse. PLoS Pathogens **2:**e11.
18. **Lee, M. S., B. A. Dougherty, A. C. Madeo, and D. A. Morrison.** 1999. Construction and analysis of a library for random insertional mutagenesis in *Streptococcus pneumoniae*: use for recovery of mutants defective in genetic transformation and for identification of essential genes. Appl. Environ. Microbiol. **65:**1883–1890.
19. **Levy, S. B., and B. Marshall.** 2004. Antibacterial resistance worldwide: causes, challenges and responses. Nat. Med. **10:**S122–S129.
20. **Long, A. D., H. J. Mangalam, B. Y. Chan, L. Tolleri, G. W. Hatfield, and P. Baldi.** 2001. Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. J. Biol. Chem. **276:**19937–19944.
21. **Maguin, E., H. Prevost, S. D. Ehrlich, and A. Gruss.** 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. J. Bacteriol. **178:**931–935.
22. **Marra, A., J. Asundi, M. Bartilson, S. Lawson, F. Fang, J. Christine, C. Wiesner, D. Brigham, W. P. Schneider, and A. E. Hromockyj.** 2002. Differential fluorescence induction analysis of *Streptococcus pneumoniae* identifies genes involved in pathogenesis. Infect. Immun. **70:**1422–1433.
23. **Martin, B., M. Prudhomme, G. Alloing, C. Granadel, and J. P. Claverys.** 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **38:**867–878.
24. **Milanino, R., M. Marrella, R. Gasperini, M. Pasqualicchio, and G. Velo.** 1993. Copper and zinc body levels in inflammation: an overview of the data obtained from animal and human studies. Agents Actions **39:**195–209.
25. **Ng, W. L., K. M. Kazmierczak, and M. E. Winkler.** 2004. Defective cell wall synthesis in *Streptococcus pneumoniae* R6 depleted for the essential PcsB putative murein hydrolase or the VicR (YycF) response regulator. Mol. Microbiol. **53:**1161–1175.
26. **Nies, D. H., and S. Silver.** 1995. Ion efflux systems involved in bacterial metal resistances. J. Ind. Microbiol. **14:**186–199.
27. **Obaro, S., and R. Adegbola.** 2002. The pneumococcus: carriage, disease and conjugate vaccines. J. Med. Microbiol. **51:**98–104.
28. **Obert, C., J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. I. Tuomanen, and C. J. Orihuela.** 2006. Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect. Immun. **74:**4766–4777.
29. **Orihuela, C. J., J. N. Radin, J. E. Sublett, G. Gao, D. Kaushal, and E. I. Tuomanen.** 2004. Microarray analysis of pneumococcal gene expression during invasive disease. Infect. Immun. **72:**5582–5596.
30. **Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari, and D. Simon.** 1998. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:**5620–5629.
31. **Rubin, E. J., B. J. Akerley, V. N. Novik, D. J. Lampe, R. N. Husson, and J. J. Mekalanos.** 1999. In vivo transposition of *mariner*-based elements in enteric bacteria and mycobacteria. Proc. Natl. Acad. Sci. USA **96:**1645–1650.
32. **Salama, N. R., B. Shepherd, and S. Falkow.** 2004. Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. J. Bacteriol. **186:**7926–7935.
33. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc. Natl. Acad. Sci. USA **98:**12712–12717.
34. **Shouval, D. S., D. Greenberg, N. Givon-Lavi, N. Porat, and R. Dagan.** 2006. Site-specific disease potential of individual *Streptococcus pneumoniae* serotypes in pediatric invasive disease, acute otitis media and acute conjunctivitis. Pediatr. Infect. Dis. J. **25:**602–607.
35. **Song, J. H., K. S. Ko, J. Y. Lee, J. Y. Baek, W. S. Oh, H. S. Yoon, J. Y. Jeong, and J. Chun.** 2005. Identification of essential genes in *Streptococcus pneumoniae* by allelic replacement mutagenesis. Mol. Cells **19:**365–374.
36. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498–506.
37. **Thanassi, J. A., S. L. Hartman-Neumann, T. J. Dougherty, B. A. Dougherty, and M. J. Pucci.** 2002. Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. Nucleic Acids Res. **30:**3152–3162.
38. **Thurnham, D. I., A. S. Mburu, D. L. Mwaniki, and A. De Wagt.** 2005. Micronutrients in childhood and the influence of subclinical inflammation. Proc. Nutr. Soc. **64:**502–509.
39. **Tong, X., J. W. Campbell, G. Balazsi, K. A. Kay, B. L. Wanner, S. Y. Gerdes, and Z. N. Oltvai.** 2004. Genome-scale identification of conditionally essential genes in *E. coli* by DNA microarrays. Biochem. Biophys. Res. Commun. **322:**347–354.
40. **van Hijum, S. A., A. de Jong, R. J. Baerends, H. A. Karsens, N. E. Kramer, R. Larsen, C. D. den Hengst, C. J. Albers, J. Kok, and O. P. Kuipers.** 2005. A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. BMC Genomics **6:**77.
41. **van Hijum, S. A., J. Garcia de la Nava, O. Trelles, J. Kok, and O. P. Kuipers.** 2003. MicroPreP: a cDNA microarray data pre-processing framework. Appl. Bioinformatics **2:**241–244.
42. **Versieck, J.** 1985. Trace elements in human body fluids and tissues. Crit. Rev. Clin. Lab. Sci. **22:**97–184.
43. **Zalacain, M., S. Biswas, K. A. Ingraham, J. Ambrad, A. Bryant, A. F. Chalker, S. Iordanescu, J. Fan, F. Fan, R. D. Lunsford, K. O'Dwyer, L. M. Palmer, C. So, D. Sylvester, C. Volker, P. Warren, D. McDevitt, J. R. Brown, D. J. Holmes, and M. K. Burnham.** 2003. A global approach to identify novel broad-spectrum antibacterial targets among proteins of unknown function. J. Mol. Microbiol. Biotechnol. **6:**109–126.